



SEQUENCE LISTING

<110> ADNEY, WILLIAM S
DING, SHI-YOU S
VINZANT, TODD B.
DECKER, STEPHEN R.
HIMMEL, MICHAEL E.

<120> THERMAL TOLERANT EXOGLUCANASE FROM ACIDOTHERMUS CELLULOLYTICUS

<130> NREL 01-38

<140> US 09/917,384
<141> 2001-07-28

<160> 11

<170> PatentIn version 3.2

<210> 1
<211> 1121
<212> PRT
<213> Acidothermus cellulolyticus

<220>
<221> misc_feature
<223> Full-length sequence of Gux1 protein

<400> 1

Met Pro Gly Leu Arg Arg Arg Leu Arg Ala Gly Ile Val Ser Ala Ala
1 5 10 15

Ala Leu Gly Ser Leu Val Ser Gly Leu Val Ala Val Ala Pro Val Ala
20 25 30

His Ala Ala Val Thr Leu Lys Ala Gln Tyr Lys Asn Asn Asp Ser Ala
35 40 45

Pro Ser Asp Asn Gln Ile Lys Pro Gly Leu Gln Leu Val Asn Thr Gly
50 55 60

Ser Ser Ser Val Asp Leu Ser Thr Val Thr Val Arg Tyr Trp Phe Thr
65 70 75 80

Arg Asp Gly Gly Ser Ser Thr Leu Val Tyr Asn Cys Asp Trp Ala Ala
85 90 95

Met Gly Cys Gly Asn Ile Arg Ala Ser Phe Gly Ser Val Asn Pro Ala
100 105 110

Thr Pro Thr Ala Asp Thr Tyr Leu Gln Leu Ser Phe Thr Gly Gly Thr
115 120 125

Leu Ala Ala Gly Gly Ser Thr Gly Glu Ile Gln Asn Arg Val Asn Lys
130 135 140

Ser Asp Trp Ser Asn Phe Asp Glu Thr Asn Asp Tyr Ser Tyr Gly Thr
145 150 155 160

Asn Thr Thr Phe Gln Asp Trp Thr Lys Val Thr Val Tyr Val Asn Gly
165 170 175

Val Leu Val Trp Gly Thr Glu Pro Ser Gly Ala Thr Ala Ser Pro Ser
180 185 190

Ala Ser Ala Thr Pro Ser Pro Ser Ser Pro Thr Thr Ser Pro Ser
195 200 205

Ser Ser Pro Ser Pro Ser Ser Ser Pro Thr Pro Thr Pro Ser Ser Ser
210 215 220

Ser Pro Pro Pro Ser Ser Asn Asp Pro Tyr Ile Gln Arg Phe Leu Thr
225 230 235 240

Met Tyr Asn Lys Ile His Asp Pro Ala Asn Gly Tyr Phe Ser Pro Gln
245 250 255

Gly Ile Pro Tyr His Ser Val Glu Thr Leu Ile Val Glu Ala Pro Asp
260 265 270

Tyr Gly His Glu Thr Thr Ser Glu Ala Tyr Ser Phe Trp Leu Trp Leu
275 280 285

Glu Ala Thr Tyr Gly Ala Val Thr Gly Asn Trp Thr Pro Phe Asn Asn
290 295 300

Ala Trp Thr Thr Met Glu Thr Tyr Met Ile Pro Gln His Ala Asp Gln
305 310 315 320

Pro Asn Asn Ala Ser Tyr Asn Pro Asn Ser Pro Ala Ser Tyr Ala Pro
325 330 335

Glu Glu Pro Leu Pro Ser Met Tyr Pro Val Ala Ile Asp Ser Ser Val
340 345 350

Pro Val Gly His Asp Pro Leu Ala Ala Glu Leu Gln Ser Thr Tyr Gly
355 360 365

Thr Pro Asp Ile Tyr Gly Met His Trp Leu Ala Asp Val Asp Asn Ile
370 375 380

Tyr Gly Tyr Gly Asp Ser Pro Gly Gly Gly Cys Glu Leu Gly Pro Ser
385 390 395 400

Ala Lys Gly Val Ser Tyr Ile Asn Thr Phe Gln Arg Gly Ser Gln Glu
 405 410 415
 Ser Val Trp Glu Thr Val Thr Gln Pro Thr Cys Asp Asn Gly Lys Tyr
 420 425 430
 Gly Gly Ala His Gly Tyr Val Asp Leu Phe Ile Gln Gly Ser Thr Pro
 435 440 445
 Pro Gln Trp Lys Tyr Thr Asp Ala Pro Asp Ala Asp Ala Arg Ala Val
 450 455 460
 Gln Ala Ala Tyr Trp Ala Tyr Thr Trp Ala Ser Ala Gln Gly Lys Ala
 465 470 475 480
 Ser Ala Ile Ala Pro Thr Ile Ala Lys Ala Ser Gln Thr Gly Asp Tyr
 485 490 495
 Leu Arg Tyr Ser Leu Phe Asp Lys Tyr Phe Lys Gln Val Gly Asn Cys
 500 505 510
 Tyr Pro Ala Ser Ser Cys Pro Gly Ala Thr Gly Arg Gln Ser Glu Thr
 515 520 525
 Tyr Leu Ile Gly Trp Tyr Tyr Ala Trp Gly Gly Ser Ser Gln Gly Trp
 530 535 540
 Ala Trp Arg Ile Gly Asp Gly Ala Ala His Phe Gly Tyr Gln Asn Pro
 545 550 555 560
 Leu Ala Ala Trp Ala Met Ser Asn Val Thr Pro Leu Ile Pro Leu Ser
 565 570 575
 Pro Thr Ala Lys Ser Asp Trp Ala Ala Ser Leu Gln Arg Gln Leu Glu
 580 585 590
 Phe Tyr Gln Trp Leu Gln Ser Ala Glu Gly Ala Ile Ala Gly Gly Ala
 595 600 605
 Thr Asn Ser Trp Asn Gly Asn Tyr Gly Thr Pro Pro Ala Gly Asp Ser
 610 615 620
 Thr Phe Tyr Gly Met Ala Tyr Asp Trp Glu Pro Val Tyr His Asp Pro
 625 630 635 640
 Pro Ser Asn Asn Trp Phe Gly Phe Gln Ala Trp Ser Met Glu Arg Val
 645 650 655

Ala Glu Tyr Tyr Tyr Val Thr Gly Asp Pro Lys Ala Lys Ala Leu Leu
 660 665 670
 Asp Lys Trp Val Ala Trp Val Lys Pro Asn Val Thr Thr Gly Ala Ser
 675 680 685
 Trp Ser Ile Pro Ser Asn Leu Ser Trp Ser Gly Gln Pro Asp Thr Trp
 690 695 700
 Asn Pro Ser Asn Pro Gly Thr Asn Ala Asn Leu His Val Thr Ile Thr
 705 710 715 720
 Ser Ser Gly Gln Asp Val Gly Val Ala Ala Ala Leu Ala Lys Thr Leu
 725 730 735
 Glu Tyr Tyr Ala Ala Lys Ser Gly Asp Thr Ala Ser Arg Asp Leu Ala
 740 745 750
 Lys Gly Leu Leu Asp Ser Met Trp Asn Asn Asp Gln Asp Ser Leu Gly
 755 760 765
 Val Ser Thr Pro Glu Thr Arg Thr Asp Tyr Ser Arg Phe Thr Gln Val
 770 775 780
 Tyr Asp Pro Thr Thr Gly Asp Gly Leu Tyr Ile Pro Ser Gly Trp Thr
 785 790 795 800
 Gly Thr Met Pro Asn Gly Asp Gln Ile Lys Pro Gly Ala Thr Phe Leu
 805 810 815
 Ser Ile Arg Ser Trp Tyr Thr Lys Asp Pro Gln Trp Ser Lys Val Gln
 820 825 830
 Ala Tyr Leu Asn Gly Gly Pro Ala Pro Thr Phe Asn Tyr His Arg Phe
 835 840 845
 Trp Ala Glu Ser Asp Phe Ala Met Ala Asn Ala Asp Phe Gly Met Leu
 850 855 860
 Phe Pro Ser Gly Ser Pro Ser Pro Thr Pro Ser Pro Thr Pro Thr Ser
 865 870 875 880
 Ser Pro Ser Pro Thr Pro Ser Ser Ser Pro Thr Pro Ser Pro Ser Pro
 885 890 895
 Ser Pro Thr Gly Asp Thr Thr Pro Pro Ser Val Pro Thr Gly Leu Gln
 900 905 910
 Val Thr Gly Thr Thr Thr Ser Ser Val Ser Leu Ser Trp Thr Ala Ser

915

920

925

Thr Asp Asn Val Gly Val Ala His Tyr Asn Val Tyr Arg Asn Gly Thr
930 935 940

Leu Val Gly Gln Pro Thr Ala Thr Ser Phe Thr Asp Thr Gly Leu Ala
945 950 955 960

Ala Gly Thr Ser Tyr Thr Tyr Thr Val Ala Ala Val Asp Ala Ala Gly
965 970 975

Asn Thr Ser Ala Gln Ser Phe Ala Gly Asp Ser Asp Asp Gly Ile Ala
980 985 990

Val Ala Ser Pro Ser Pro Ser Pro Thr Pro Thr Ser Ser Pro Ser Pro
995 1000 1005

Thr Pro Ser Pro Thr Pro Ser Pro Thr Ser Thr Ser Gly Ala Ser
1010 1015 1020

Cys Thr Ala Thr Tyr Val Val Asn Ser Asp Trp Gly Ser Gly Phe
1025 1030 1035

Thr Thr Thr Val Thr Val Thr Asn Thr Gly Thr Arg Ala Thr Ser
1040 1045 1050

Gly Trp Thr Val Thr Trp Ser Phe Ala Gly Asn Gln Thr Val Thr
1055 1060 1065

Asn Tyr Trp Asn Thr Ala Leu Thr Gln Ser Gly Lys Ser Val Thr
1070 1075 1080

Ala Lys Asn Leu Ser Tyr Asn Asn Val Ile Gln Pro Gly Gln Ser
1085 1090 1095

Thr Thr Phe Gly Phe Asn Gly Ser Tyr Ser Gly Thr Asn Thr Ala
1100 1105 1110

Pro Thr Leu Ser Cys Thr Ala Ser
1115 1120

<210> 2
<211> 3365
<212> DNA
<213> Acidothermus cellulolyticus

<220>
<221> misc_feature
<223> Gux1 full-length coding sequence

<400> 2

atgccaggat tacgacggcg actccgcgcc ggtatcgtct cggcggcggc gttggggctcg	60
ctggtttagcg ggctcgttgc cgtcgcacca gtcgcgcacg cggcgggtgac tctcaaagcg	120
cagtataaga acaatgattc ggcgccgagt gacaaccaga tcaaaccggg tctccagttg	180
gtgaataaccg ggtcgtcgtc ggtggatttg tcgacgggtga cggtgcggtta ctggttcacc	240
cgggatggtg ggtcgtcgc actggtgtac aactgtgact gggcggcgat ggggtgtggg	300
aatatccgcg cctcgttcgg ctcggtgaac cggcgacgc cgacggcgga cacctacctg	360
cagttgtcgt tctactggtg aacgttgcc gctggtgggt cgacgggtga gattcaaaac	420
cgggtgaata agagtgactg gtcgaacttt gatgagacca atgactactc gtatgggacg	480
aacaccacct tccaggactg gacgaagggt acggtgtacg tcaacggcgt gttggtctgg	540
gggaccgaac cgtccggagc gacggcgtct ccatccgcgt cggcgacgcc cagcccgtcc	600
agttcaccga ccacgagtcc gagttcgtcc ccgtcgccga gcagcagccc gacgccgaca	660
ccgagcagct cgtcgccgcc ccgtcgtcca acgaccgta catccagcgg ttcctcacga	720
tgtacaacaa gattcacgac ccagcgaacg gctacttcag cccgcaggga attccctacc	780
actcggtaga aacgctcatc gttgaggcac cggactacgg gcacgagaca acttcggagg	840
cgtacagctt ctggctctgg ctggaagcga cgtacggcgc agtgaccggc aactggacgc	900
cgttcaacaa cgcctggacg acgatggaaa cgtacatgat cccgcagcac gcggaccagc	960
cgaacaacgc gtcgtacaac cccaacagcc cggcgtcgta cgctccggaa gagccgctgc	1020
ccagcatgta cccggttgcc atcgacagca gcgtgccggg tgggcacgac ccgctcgccg	1080
ccgaattgca gtcgacgtac ggcaactccg acatttacgg catgcaactg ctggccgacg	1140
ttgacaacat ctacggatac ggcgacagcc ccggcgggtg ttgcgaactc ggtccttccg	1200
ctaagggcgt ctctacatc aacacattcc agcgcggctc gcaggagtcc gtctgggaga	1260
cggtcacca gccgacgtgc gacaacggca agtacggtgg ggcgcacggc tacgtcgacc	1320
tgttcatcca gggttcgacg ccgccgcagt ggaagtacac cgatgccccg gacgccgacg	1380
cccgtgccgt ccaggctgcg tactgggcct acacctgggc atcggcgagc ggcaaggcaa	1440
gcgcgattgc cccgacgatc gccaaaggcga gccaaaccgg cgactacctg cggactcgc	1500
tctttgacaa gtacttcaag caggtcggca actgctaccc ggccagctcc tgccctggag	1560
caaccggacg ccagagcgag acctacctga tcggctggta ctacgcctgg ggcggctcaa	1620
gccaaaggctg ggcctggcgc attggtgacg gcgccgcgca cttcggctac cagaatccgc	1680
ttgccgcgtg ggcgatgtcg aacgtgacac cgctcattcc gctctcgccc acggcaaaga	1740
gcgactgggc ggcgagcttg cagcgccagc tggagttcta ccagtggttg caatccgcgg	1800
aaggagccat tgcggggcgg gccaccaaca gctggaacgg caattacggg accccgcccg	1860
ccggagactc gaccttctac ggcattggcgt acgactggga gccggtctac cagacccgc	1920

cgagcaacaa ctggttcggc ttccaggcgt ggtccatgga acgggttgcc gagtactact	1980
acgtcaccgg cgacccgaag gccaaggcgc tgctcgacaa gtgggtcgca tgggtgaagc	2040
cgaatgtcac caccggtgcc tcatggtcga ttccgtcgaa tttgtcctgg agcggccaac	2100
cggatacctg gaatccgagc aaccaggaa cgaatgccaa cctgcacgtg accatcacgt	2160
cgtccgggca ggacgtcggg gttgccgcgg cgctcgcgaa gacactcgag tactacgcgg	2220
caaaatccgg cgatacggcc tcgcgcgacc tcgcgaaggg attgctcgac tccatgtgga	2280
acaacgacca ggacagcctc ggtgtgagca caccggagac gcggaccgac tactctcggg	2340
tcactcaggt gtacgacccg acgactgggt acggcctcta catcccgtcg ggttggacgg	2400
ggaccatgcc caacggtgac caaatcaagc cgggtgcgac cttcctgagc atccggtcct	2460
ggtacaccaa ggatccgcag tggtcgaagg tgcaggcgta cctcaacggc gggcctgctc	2520
cgacgttcaa ctaccaccgg ttctgggcgg agtccgactt cgcgatggcg aacgccgatt	2580
ttggcatgct cttcccatcc gggtcgcccc gcccgacccc gagccccgact ccgacgtcgt	2640
ccccgagccc gactccgagc agctcgccga cgccgtcgcc cagcccgtca ccgaccggcg	2700
acaccacgcc gccgagcgtg ccgacgggtc ttcaggtcac cgggacaacg acgtcgtccg	2760
tgctgctcag ctggaccgcg tccaccgaca acgtcggcgt cgcgactac aacgtgtacc	2820
gaaacggcac gctggtgggt cagccgacag cgacgtcgtt cacggacacc ggcctggctg	2880
ctggcacgtc gtacacgtac acagtggcgg ccgttgatgc ggccggtaac acgtcggcgc	2940
agagcttcgc cggtgacagc gacgacggca tcgccgtcgc gagcccgtcg ccgagcccga	3000
ctccgacgtc gtccccgagc ccaacgccgt cgccgacacc gtcaccgacg tccaccagcg	3060
gcgcacgtg cactgctacc tacgttgtca atagcgactg gggtagcggc ttcacgacaa	3120
ccgtgaccgt gacgaacacc ggcaccaggg ccaccagtgg ctggacggtc acgtggagct	3180
ttgccggtaa tcagacggtc accaactact ggaacaccgc gctgacgcaa tccggaaagt	3240
cggtgaccgc aaagaacctg agttacaaca acgtcatcca acctggtcag tcgacgacct	3300
ttggattcaa cggaagttac tcaggaacaa acaccgcgcc gacgctcagc tgcacggcaa	3360
gctga	3365

<210> 3
 <211> 34
 <212> PRT
 <213> Acidothermus cellulolyticus

<220>
 <221> misc_feature
 <223> Potential signal peptide of Gux1

<400> 3

Met	Pro	Gly	Leu	Arg	Arg	Arg	Leu	Arg	Ala	Gly	Ile	Val	Ser	Ala	Ala
1				5					10					15	

Ala Leu Gly Ser Leu Val Ser Gly Leu Val Ala Val Ala Pro Val Ala
20 25 30

His Ala

<210> 4
<211> 153
<212> PRT
<213> Acidothermus cellulolyticus

<220>
<221> misc_feature
<223> CBD III of Gux1

<400> 4

Ala Val Thr Leu Lys Ala Gln Tyr Lys Asn Asn Asp Ser Ala Pro Ser
1 5 10 15

Asp Asn Gln Ile Lys Pro Gly Leu Gln Leu Val Asn Thr Gly Ser Ser
20 25 30

Ser Val Asp Leu Ser Thr Val Thr Val Arg Tyr Trp Phe Thr Arg Asp
35 40 45

Gly Gly Ser Ser Thr Leu Val Tyr Asn Cys Asp Trp Ala Ala Met Gly
50 55 60

Cys Gly Asn Ile Arg Ala Ser Phe Gly Ser Val Asn Pro Ala Thr Pro
65 70 75 80

Thr Ala Asp Thr Tyr Leu Gln Leu Ser Phe Thr Gly Gly Thr Leu Ala
85 90 95

Ala Gly Gly Ser Thr Gly Glu Ile Gln Asn Arg Val Asn Lys Ser Asp
100 105 110

Trp Ser Asn Phe Asp Glu Thr Asn Asp Tyr Ser Tyr Gly Thr Asn Thr
115 120 125

Thr Phe Gln Asp Trp Thr Lys Val Thr Val Tyr Val Asn Gly Val Leu
130 135 140

Val Trp Gly Thr Glu Pro Ser Gly Ala
145 150

<210> 5
<211> 640
<212> PRT

<213> Acidothermus cellulolyticus

<220>

<221> misc_feature

<223> GH48 domain of Gux1

<400> 5

Asn Asp Pro Tyr Ile Gln Arg Phe Leu Thr Met Tyr Asn Lys Ile His
1 5 10 15

Asp Pro Ala Asn Gly Tyr Phe Ser Pro Gln Gly Ile Pro Tyr His Ser
20 25 30

Val Glu Thr Leu Ile Val Glu Ala Pro Asp Tyr Gly His Glu Thr Thr
35 40 45

Ser Glu Ala Tyr Ser Phe Trp Leu Trp Leu Glu Ala Thr Tyr Gly Ala
50 55 60

Val Thr Gly Asn Trp Thr Pro Phe Asn Asn Ala Trp Thr Thr Met Glu
65 70 75 80

Thr Tyr Met Ile Pro Gln His Ala Asp Gln Pro Asn Asn Ala Ser Tyr
85 90 95

Asn Pro Asn Ser Pro Ala Ser Tyr Ala Pro Glu Glu Pro Leu Pro Ser
100 105 110

Met Tyr Pro Val Ala Ile Asp Ser Ser Val Pro Val Gly His Asp Pro
115 120 125

Leu Ala Ala Glu Leu Gln Ser Thr Tyr Gly Thr Pro Asp Ile Tyr Gly
130 135 140

Met His Trp Leu Ala Asp Val Asp Asn Ile Tyr Gly Tyr Gly Asp Ser
145 150 155 160

Pro Gly Gly Gly Cys Glu Leu Gly Pro Ser Ala Lys Gly Val Ser Tyr
165 170 175

Ile Asn Thr Phe Gln Arg Gly Ser Gln Glu Ser Val Trp Glu Thr Val
180 185 190

Thr Gln Pro Thr Cys Asp Asn Gly Lys Tyr Gly Gly Ala His Gly Tyr
195 200 205

Val Asp Leu Phe Ile Gln Gly Ser Thr Pro Pro Gln Trp Lys Tyr Thr
210 215 220

Asp Ala Pro Asp Ala Asp Ala Arg Ala Val Gln Ala Ala Tyr Trp Ala
 225 230 235 240
 Tyr Thr Trp Ala Ser Ala Gln Gly Lys Ala Ser Ala Ile Ala Pro Thr
 245 250 255
 Ile Ala Lys Ala Ser Gln Thr Gly Asp Tyr Leu Arg Tyr Ser Leu Phe
 260 265 270
 Asp Lys Tyr Phe Lys Gln Val Gly Asn Cys Tyr Pro Ala Ser Ser Cys
 275 280 285
 Pro Gly Ala Thr Gly Arg Gln Ser Glu Thr Tyr Leu Ile Gly Trp Tyr
 290 295 300
 Tyr Ala Trp Gly Gly Ser Ser Gln Gly Trp Ala Trp Arg Ile Gly Asp
 305 310 315 320
 Gly Ala Ala His Phe Gly Tyr Gln Asn Pro Leu Ala Ala Trp Ala Met
 325 330 335
 Ser Asn Val Thr Pro Leu Ile Pro Leu Ser Pro Thr Ala Lys Ser Asp
 340 345 350
 Trp Ala Ala Ser Leu Gln Arg Gln Leu Glu Phe Tyr Gln Trp Leu Gln
 355 360 365
 Ser Ala Glu Gly Ala Ile Ala Gly Gly Ala Thr Asn Ser Trp Asn Gly
 370 375 380
 Asn Tyr Gly Thr Pro Pro Ala Gly Asp Ser Thr Phe Tyr Gly Met Ala
 385 390 395 400
 Tyr Asp Trp Glu Pro Val Tyr His Asp Pro Pro Ser Asn Asn Trp Phe
 405 410 415
 Gly Phe Gln Ala Trp Ser Met Glu Arg Val Ala Glu Tyr Tyr Tyr Val
 420 425 430
 Thr Gly Asp Pro Lys Ala Lys Ala Leu Leu Asp Lys Trp Val Ala Trp
 435 440 445
 Val Lys Pro Asn Val Thr Thr Gly Ala Ser Trp Ser Ile Pro Ser Asn
 450 455 460
 Leu Ser Trp Ser Gly Gln Pro Asp Thr Trp Asn Pro Ser Asn Pro Gly
 465 470 475 480
 Thr Asn Ala Asn Leu His Val Thr Ile Thr Ser Ser Gly Gln Asp Val

485

490

495

Gly Val Ala Ala Ala Leu Ala Lys Thr Leu Glu Tyr Tyr Ala Ala Lys
500 505 510

Ser Gly Asp Thr Ala Ser Arg Asp Leu Ala Lys Gly Leu Leu Asp Ser
515 520 525

Met Trp Asn Asn Asp Gln Asp Ser Leu Gly Val Ser Thr Pro Glu Thr
530 535 540

Arg Thr Asp Tyr Ser Arg Phe Thr Gln Val Tyr Asp Pro Thr Thr Gly
545 550 555 560

Asp Gly Leu Tyr Ile Pro Ser Gly Trp Thr Gly Thr Met Pro Asn Gly
565 570 575

Asp Gln Ile Lys Pro Gly Ala Thr Phe Leu Ser Ile Arg Ser Trp Tyr
580 585 590

Thr Lys Asp Pro Gln Trp Ser Lys Val Gln Ala Tyr Leu Asn Gly Gly
595 600 605

Pro Ala Pro Thr Phe Asn Tyr His Arg Phe Trp Ala Glu Ser Asp Phe
610 615 620

Ala Met Ala Asn Ala Asp Phe Gly Met Leu Phe Pro Ser Gly Ser Pro
625 630 635 640

```
<210> 6
<211> 85
<212> PRT
<213> Acidothermus cellulolyticus
```

```
<220>
<221> misc_feature
<223> FN III domain of Gux1
```

<400> 6

Asp Thr Thr Pro Pro Ser Val Pro Thr Gly Leu Gln Val Thr Gly Thr
1 5 10 15

Thr Thr Ser Ser Val Ser Leu Ser Trp Thr Ala Ser Thr Asp Asn Val
20 25 30

Gly Val Ala His Tyr Asn Val Tyr Arg Asn Gly Thr Leu Val Gly Gln
35 40 45

Pro Thr Ala Thr Ser Phe Thr Asp Thr Gly Leu Ala Ala Gly Thr Ser
50 55 60

Tyr Thr Tyr Thr Val Ala Ala Val Asp Ala Ala Gly Asn Thr Ser Ala
65 70 75 80

Gln Ser Phe Ala Gly
85

<210> 7
<211> 101
<212> PRT
<213> *Acidothermus cellulolyticus*

<220>
<221> misc_feature
<223> CBD II domain of Gux1

<400> 7

Gly Ala Ser Cys Thr Ala Thr Tyr Val Val Asn Ser Asp Trp Gly Ser
1 5 10 15

Gly Phe Thr Thr Thr Val Thr Val Thr Asn Thr Gly Thr Arg Ala Thr
20 25 30

Ser Gly Trp Thr Val Thr Trp Ser Phe Ala Gly Asn Gln Thr Val Thr
35 40 45

Asn Tyr Trp Asn Thr Ala Leu Thr Gln Ser Gly Lys Ser Val Thr Ala
50 55 60

Lys Asn Leu Ser Tyr Asn Asn Val Ile Gln Pro Gly Gln Ser Thr Thr
65 70 75 80

Phe Gly Phe Asn Gly Ser Tyr Ser Gly Thr Asn Thr Ala Pro Thr Leu
85 90 95

Ser Cys Thr Ala Ser
100

<210> 8
<211> 6
<212> PRT
<213> Artificial

<220>
<223> synthetic histidine tag

<220>
<221> misc_feature
<223> Synthetic His Tag

<400> 8

His His His His His His
1 5

<210> 9
<211> 638
<212> PRT
<213> Acidothermus cellulolyticus

<220>
<221> misc_feature
<223> GH48 domain of Acidothermus cellulolyticus

<400> 9

Pro Tyr Ile Gln Arg Phe Leu Thr Met Tyr Asn Lys Ile His Asp Pro
1 5 10 15

Ala Asn Gly Tyr Phe Ser Pro Gln Gly Ile Pro Tyr His Ser Val Glu
20 25 30

Thr Leu Ile Val Glu Ala Pro Asp Tyr Gly His Glu Thr Thr Ser Glu
35 40 45

Ala Tyr Ser Phe Trp Leu Trp Leu Glu Ala Thr Tyr Gly Ala Val Thr
50 55 60

Gly Asn Trp Thr Pro Phe Asn Asn Ala Trp Thr Thr Met Glu Thr Tyr
65 70 75 80

Met Ile Pro Gln His Ala Asp Gln Pro Asn Asn Ala Ser Tyr Asn Pro
85 90 95

Asn Ser Pro Ala Ser Tyr Ala Pro Glu Glu Pro Leu Pro Ser Met Tyr
100 105 110

Pro Val Ala Ile Asp Ser Ser Val Pro Val Gly His Asp Pro Leu Ala
115 120 125

Ala Glu Leu Gln Ser Thr Tyr Gly Thr Pro Asp Ile Tyr Gly Met His
130 135 140

Trp Leu Ala Asp Val Asp Asn Ile Tyr Gly Tyr Gly Asp Ser Pro Gly
145 150 155 160

Gly Gly Cys Glu Leu Gly Pro Ser Ala Lys Gly Val Ser Tyr Ile Asn
165 170 175

Thr Phe Gln Arg Gly Ser Gln Glu Ser Val Trp Glu Thr Val Thr Gln
180 185 190

Pro Thr Cys Asp Asn Gly Lys Tyr Gly Gly Ala His Gly Tyr Val Asp

195

200

205

Leu Phe Ile Gln Gly Ser Thr Pro Pro Gln Trp Lys Tyr Thr Asp Ala
 210 215 220

Pro Asp Ala Asp Ala Arg Ala Val Gln Ala Ala Tyr Trp Ala Tyr Thr
 225 230 235 240

Trp Ala Ser Ala Gln Gly Lys Ala Ser Ala Ile Ala Pro Thr Ile Ala
 245 250 255

Lys Ala Ser Gln Thr Gly Asp Tyr Leu Arg Tyr Ser Leu Phe Asp Lys
 260 265 270

Tyr Phe Lys Gln Val Gly Asn Cys Tyr Pro Ala Ser Ser Cys Pro Gly
 275 280 285

Ala Thr Gly Arg Gln Ser Glu Thr Tyr Leu Ile Gly Trp Tyr Tyr Ala
 290 295 300

Trp Gly Gly Ser Ser Gln Gly Trp Ala Trp Arg Ile Gly Asp Gly Ala
 305 310 315 320

Ala His Phe Gly Tyr Gln Asn Pro Leu Ala Ala Trp Ala Met Ser Asn
 325 330 335

Val Thr Pro Leu Ile Pro Leu Ser Pro Thr Ala Lys Ser Asp Trp Ala
 340 345 350

Ala Ser Leu Gln Arg Gln Leu Glu Phe Tyr Gln Trp Leu Gln Ser Ala
 355 360 365

Glu Gly Ala Ile Ala Gly Gly Ala Thr Asn Ser Trp Asn Gly Asn Tyr
 370 375 380

Gly Thr Pro Pro Ala Gly Asp Ser Thr Phe Tyr Gly Met Ala Tyr Asp
 385 390 395 400

Trp Glu Pro Val Tyr His Asp Pro Pro Ser Asn Asn Trp Phe Gly Phe
 405 410 415

Gln Ala Trp Ser Met Glu Arg Val Ala Glu Tyr Tyr Tyr Val Thr Gly
 420 425 430

Asp Pro Lys Ala Lys Ala Leu Leu Asp Lys Trp Val Ala Trp Val Lys
 435 440 445

Pro Asn Val Thr Thr Gly Ala Ser Trp Ser Ile Pro Ser Asn Leu Ser
 450 455 460

Trp Ser Gly Gln Pro Asp Thr Trp Asn Pro Ser Asn Pro Gly Thr Asn
465 470 475 480

Ala Asn Leu His Val Thr Ile Thr Ser Ser Gly Gln Asp Val Gly Val
485 490 495

Ala Ala Ala Leu Ala Lys Thr Leu Glu Tyr Tyr Ala Ala Lys Ser Gly
500 505 510

Asp Thr Ala Ser Arg Asp Leu Ala Lys Gly Leu Leu Asp Ser Met Trp
515 520 525

Asn Asn Asp Gln Asp Ser Leu Gly Val Ser Thr Pro Glu Thr Arg Thr
530 535 540

Asp Tyr Ser Arg Phe Thr Gln Val Tyr Asp Pro Thr Thr Gly Asp Gly
545 550 555 560

Leu Tyr Ile Pro Ser Gly Trp Thr Gly Thr Met Pro Asn Gly Asp Gln
565 570 575

Ile Lys Pro Gly Ala Thr Phe Leu Ser Ile Arg Ser Trp Tyr Thr Lys
580 585 590

Asp Pro Gln Trp Ser Lys Val Gln Ala Tyr Leu Asn Gly Gly Pro Ala
595 600 605

Pro Thr Phe Asn Tyr His Arg Phe Trp Ala Glu Ser Asp Phe Ala Met
610 615 620

Ala Asn Ala Asp Phe Gly Met Leu Phe Pro Ser Gly Ser Pro
625 630 635

<210> 10
<211> 640
<212> PRT
<213> Cellulomonas fimi

<220>
<221> misc_feature
<223> Cellulomonas fimi CBHB

<400> 10

Glu Tyr Ala Gln Arg Phe Leu Ala Gln Tyr Asp Lys Ile Lys Asp Pro
1 5 10 15

Ala Asn Gly Tyr Phe Ser Ala Gln Gly Ile Pro Tyr His Ala Val Glu
20 25 30

Thr Leu Met Val Glu Ala Pro Asp Tyr Gly His Glu Thr Thr Ser Glu
35 40 45

Ala Tyr Ser Tyr Trp Leu Trp Leu Glu Ala Leu Tyr Gly Gln Val Thr
50 55 60

Gln Asp Trp Ala Pro Leu Asn His Ala Trp Asp Thr Met Glu Lys Tyr
65 70 75 80

Met Ile Pro Gln Ser Val Asp Gln Pro Thr Asn Ser Phe Tyr Asn Pro
85 90 95

Asn Ser Pro Ala Thr Tyr Ala Pro Glu Phe Asn His Pro Ser Ser Tyr
100 105 110

Pro Ser Gln Leu Asn Ser Gly Ile Ser Gly Gly Thr Asp Pro Ile Gly
115 120 125

Ala Glu Leu Lys Ala Thr Tyr Gly Asn Ala Asp Val Tyr Gln Met His
130 135 140

Trp Leu Ala Asp Val Asp Asn Ile Tyr Gly Phe Gly Ala Thr Pro Gly
145 150 155 160

Ala Gly Cys Thr Leu Gly Pro Thr Ala Thr Gly Thr Ser Phe Ile Asn
165 170 175

Thr Phe Gln Arg Gly Pro Gln Glu Ser Val Trp Glu Thr Val Pro Gln
180 185 190

Pro Ser Cys Glu Glu Phe Lys Tyr Gly Gly Lys Asn Gly Tyr Leu Asp
195 200 205

Leu Phe Thr Lys Asp Ala Ser Tyr Ala Lys Gln Trp Lys Tyr Thr Ser
210 215 220

Ala Ser Asp Ala Asp Ala Arg Ala Val Glu Ala Val Tyr Trp Ala Asn
225 230 235 240

Gln Trp Ala Thr Glu Gln Gly Lys Ala Ala Asp Val Ala Ala Thr Val
245 250 255

Ala Lys Ala Ala Lys Met Gly Asp Tyr Leu Arg Tyr Thr Leu Phe Asp
260 265 270

Lys Tyr Phe Lys Lys Ile Gly Cys Thr Ser Pro Thr Cys Ala Ala Gly
275 280 285

Gln Gly Arg Glu Ala Ala His Tyr Leu Leu Ser Trp Tyr Met Ala Trp
 290 300
 Gly Gly Ala Thr Asp Thr Ser Ser Gly Trp Ala Trp Arg Ile Gly Ser
 305 310 315 320
 Ser His Ala His Phe Gly Tyr Gln Asn Pro Leu Ala Ala Trp Ala Leu
 325 330 335
 Ser Thr Asp Pro Lys Leu Thr Pro Lys Ser Pro Thr Ala Lys Ala Asp
 340 345 350
 Trp Ala Ala Ser Met Gln Arg Gln Leu Glu Phe Tyr Thr Trp Leu Gln
 355 360 365
 Ala Ser Asn Gly Gly Ile Ala Gly Gly Ala Thr Asn Ser Trp Asp Gly
 370 375 380
 Ala Tyr Ala Gln Pro Pro Ala Gly Thr Pro Thr Phe Tyr Gly Met Gly
 385 390 395 400
 Tyr Thr Glu Ala Pro Val Tyr Val Asp Pro Pro Ser Asn Arg Trp Phe
 405 410 415
 Gly Met Gln Ala Trp Gly Val Gln Arg Val Ala Glu Leu Tyr Tyr Ala
 420 425 430
 Ser Gly Asn Ala Gln Ala Lys Lys Ile Leu Asp Lys Trp Val Pro Trp
 435 440 445
 Val Val Ala Asn Ile Ser Thr Asp Gly Ala Ser Trp Lys Val Pro Ser
 450 455 460
 Glu Leu Lys Trp Thr Gly Lys Pro Asp Thr Trp Asn Ala Ala Ala Pro
 465 470 475 480
 Thr Gly Asn Pro Gly Leu Thr Val Glu Val Thr Ser Tyr Gly Gln Asp
 485 490 495
 Val Gly Val Ala Ala Asp Thr Ala Arg Ala Leu Leu Phe Tyr Ala Ala
 500 505 510
 Lys Ser Gly Asp Thr Ala Ser Arg Asp Lys Ala Lys Ala Leu Leu Asp
 515 520 525
 Ala Ile Trp Ala Asn Asn Gln Asp Pro Leu Gly Val Ser Ala Val Glu
 530 535 540
 Thr Arg Gly Asp Tyr Lys Arg Phe Asp Asp Thr Tyr Val Ala Asn Gly

545 550 555 560

Asp Gly Ile Tyr Ile Pro Ser Gly Trp Thr Gly Thr Met Pro Asn Gly
565 570 575

Asp Val Ile Lys Pro Gly Val Ser Phe Leu Asp Ile Arg Ser Phe Tyr
580 585 590

Lys Lys Asp Pro Asn Trp Ser Lys Val Gln Thr Phe Leu Asp Gly Gly
595 600 605

Ala Glu Pro Gln Phe Arg Tyr His Arg Phe Trp Ala Gln Thr Ala Val
610 615 620

Ala Gly Ala Leu Ala Asp Tyr Ala Arg Leu Phe Asp Asp Gly Thr Thr
625 630 635 640

<210> 11
<211> 642
<212> PRT
<213> Thermobifida fusca

<220>
<221> misc_feature
<223> Thermobifida fusca E6

<400> 11

Ser Tyr Asp Gln Ala Phe Leu Glu Gln Tyr Glu Lys Ile Lys Asp Pro
1 5 10 15

Ala Ser Gly Tyr Phe Arg Glu Phe Asn Gly Leu Leu Val Pro Tyr His
20 25 30

Ser Val Glu Thr Met Ile Val Glu Ala Pro Asp His Gly His Gln Thr
35 40 45

Thr Ser Glu Ala Phe Ser Tyr Tyr Leu Trp Leu Glu Ala Tyr Tyr Gly
50 55 60

Arg Val Thr Gly Asp Trp Lys Pro Leu His Asp Ala Trp Glu Ser Met
65 70 75 80

Glu Thr Phe Ile Ile Pro Gly Thr Lys Asp Gln Pro Thr Asn Ser Ala
85 90 95

Tyr Asn Pro Asn Ser Pro Ala Thr Tyr Ile Pro Glu Gln Pro Asn Ala
100 105 110

Asp Gly Tyr Pro Ser Pro Leu Met Asn Asn Val Pro Val Gly Gln Asp
115 120 125

Pro Leu Ala Gln Glu Leu Ser Ser Thr Tyr Gly Thr Asn Glu Ile Tyr
130 135 140

Gly Met His Trp Leu Leu Asp Val Asp Asn Val Tyr Gly Phe Gly Phe
145 150 155 160

Cys Gly Asp Gly Thr Asp Asp Ala Pro Ala Tyr Ile Asn Thr Tyr Gln
165 170 175

Arg Gly Ala Arg Glu Ser Val Trp Glu Thr Ile Pro His Pro Ser Cys
180 185 190

Asp Asp Phe Thr His Gly Gly Pro Asn Gly Tyr Leu Asp Leu Phe Thr
195 200 205

Asp Asp Gln Asn Tyr Ala Lys Gln Trp Arg Tyr Thr Asn Ala Pro Asp
210 215 220

Ala Asp Ala Arg Ala Val Gln Val Met Phe Trp Ala His Glu Trp Ala
225 230 235 240

Lys Glu Gln Gly Lys Glu Asn Glu Ile Ala Gly Leu Met Asp Lys Ala
245 250 255

Ser Lys Met Gly Asp Tyr Leu Arg Tyr Ala Met Phe Asp Lys Tyr Phe
260 265 270

Lys Lys Ile Gly Asn Cys Val Gly Ala Thr Ser Cys Pro Gly Gly Gln
275 280 285

Gly Lys Asp Ser Ala His Tyr Leu Leu Ser Trp Tyr Tyr Ser Trp Gly
290 295 300

Gly Ser Leu Asp Thr Ser Ser Ala Trp Ala Trp Arg Ile Gly Ser Ser
305 310 315 320

Ser Ser His Gln Gly Tyr Gln Asn Val Leu Ala Ala Tyr Ala Leu Ser
325 330 335

Gln Val Pro Glu Leu Gln Pro Asp Ser Pro Thr Gly Val Gln Asp Trp
340 345 350

Ala Thr Ser Phe Asp Arg Gln Leu Glu Phe Leu Gln Trp Leu Gln Ser
355 360 365

Ala Glu Gly Gly Ile Ala Gly Gly Ala Thr Asn Ser Trp Lys Gly Ser
370 375 380

Tyr Asp Thr Pro Pro Thr Gly Leu Ser Gln Phe Tyr Gly Met Tyr Tyr
385 390 395 400

Asp Trp Gln Pro Val Trp Asn Asp Pro Pro Ser Asn Asn Trp Phe Gly
405 410 415

Phe Gln Val Trp Asn Met Glu Arg Val Ala Gln Leu Tyr Tyr Val Thr
420 425 430

Gly Asp Ala Arg Ala Glu Ala Ile Leu Asp Lys Trp Val Pro Trp Ala
435 440 445

Ile Gln His Thr Asp Val Asp Ala Asp Asn Gly Gly Gln Asn Phe Gln
450 455 460

Val Pro Ser Asp Leu Glu Trp Ser Gly Gln Pro Asp Thr Trp Thr Gly
465 470 475 480

Thr Tyr Thr Gly Asn Pro Asn Leu His Val Gln Val Val Ser Tyr Ser
485 490 495

Gln Asp Val Gly Val Thr Ala Ala Leu Ala Lys Thr Leu Met Tyr Tyr
500 505 510

Ala Lys Arg Ser Gly Asp Thr Thr Ala Leu Ala Thr Ala Glu Gly Leu
515 520 525

Leu Asp Ala Leu Leu Ala His Arg Asp Ser Ile Gly Ile Ala Thr Pro
530 535 540

Glu Gln Pro Ser Trp Asp Arg Leu Asp Asp Pro Trp Asp Gly Ser Glu
545 550 555 560

Gly Leu Tyr Val Pro Pro Gly Trp Ser Gly Thr Met Pro Asn Gly Asp
565 570 575

Arg Ile Glu Pro Gly Ala Thr Phe Leu Ser Ile Arg Ser Phe Tyr Lys
580 585 590

Asn Asp Pro Leu Trp Pro Gln Val Glu Ala His Leu Asn Asp Pro Gln
595 600 605

Asn Val Pro Ala Pro Ile Val Glu Arg His Arg Phe Trp Ala Gln Val
610 615 620

Glu Ile Ala Thr Ala Phe Ala Ala His Asp Glu Leu Phe Gly Ala Gly
625 630 635 640

Ala Pro